ITMF, ICCTM
Stickiness Working group

Measurements based on counts : variability and methods of analysis

Gozé E., Gourlot J.-P., Lassus S.

Annual cropping systems research unit, CIRAD

Bremen, March 2010

Measurements based on counts : variability and methods of analysis

Useful measurement must be

1 correlated to practical properties of the analyzed material

2 reproducible

Measurements are variable
Variability => risk in decisions
need of criteria to measure variability : e.g. CV=5%

Is a single figure of standard deviation or a CV useful for count data? Is a 5% CV a good benchmark ?

Outline

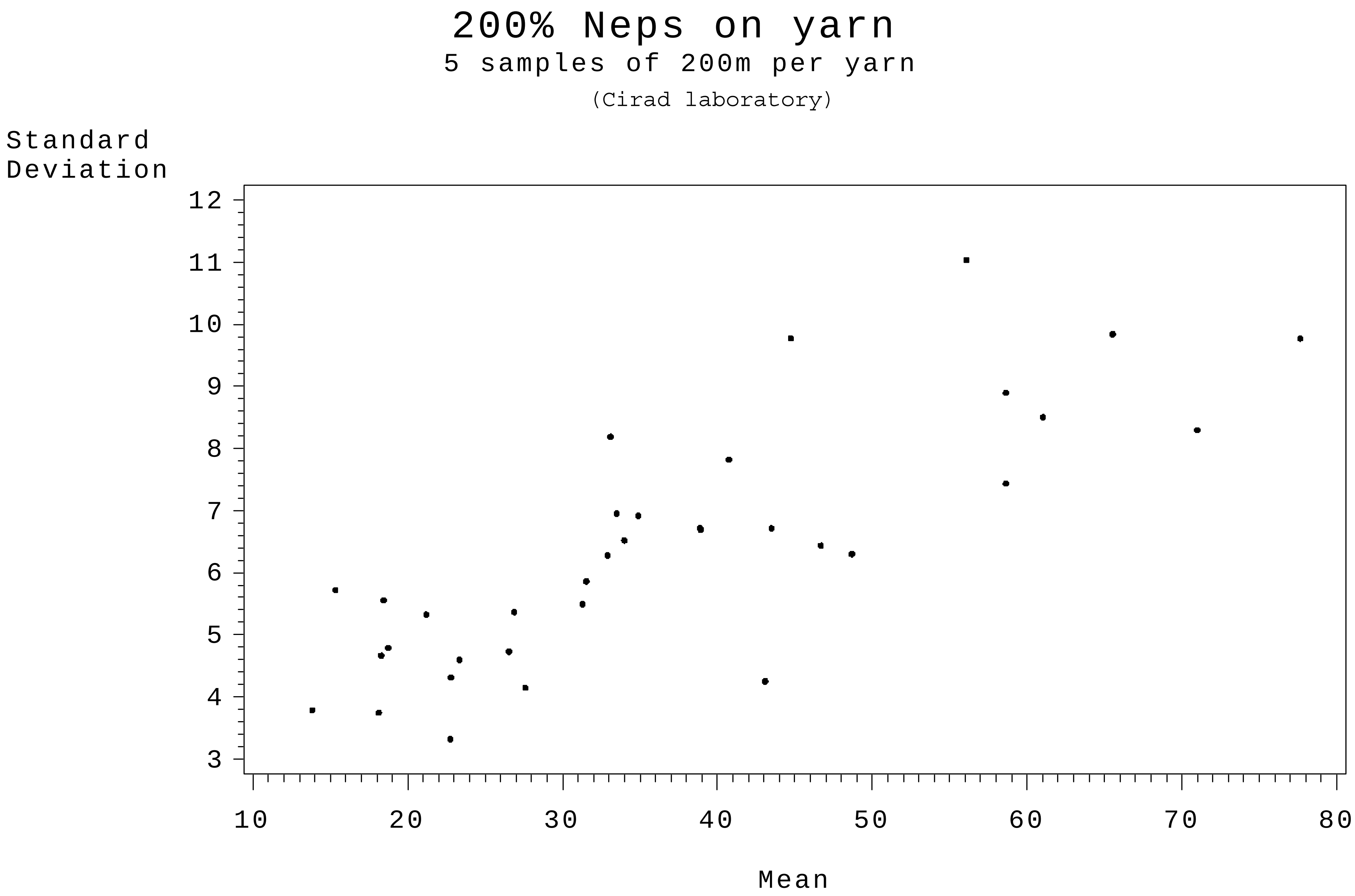
- Variability of some defects counts on yarn and fiber : evidence and characterization of a mean to variance relationship
- How to analyze calibration and round test experiments
- Research directions : extra sources of variability + confidence intervals

Outline

- Variability of some defects counts on yarn and fiber : evidence and characterization of a mean to variance relationship

- How to analyze calibration and round test experiments

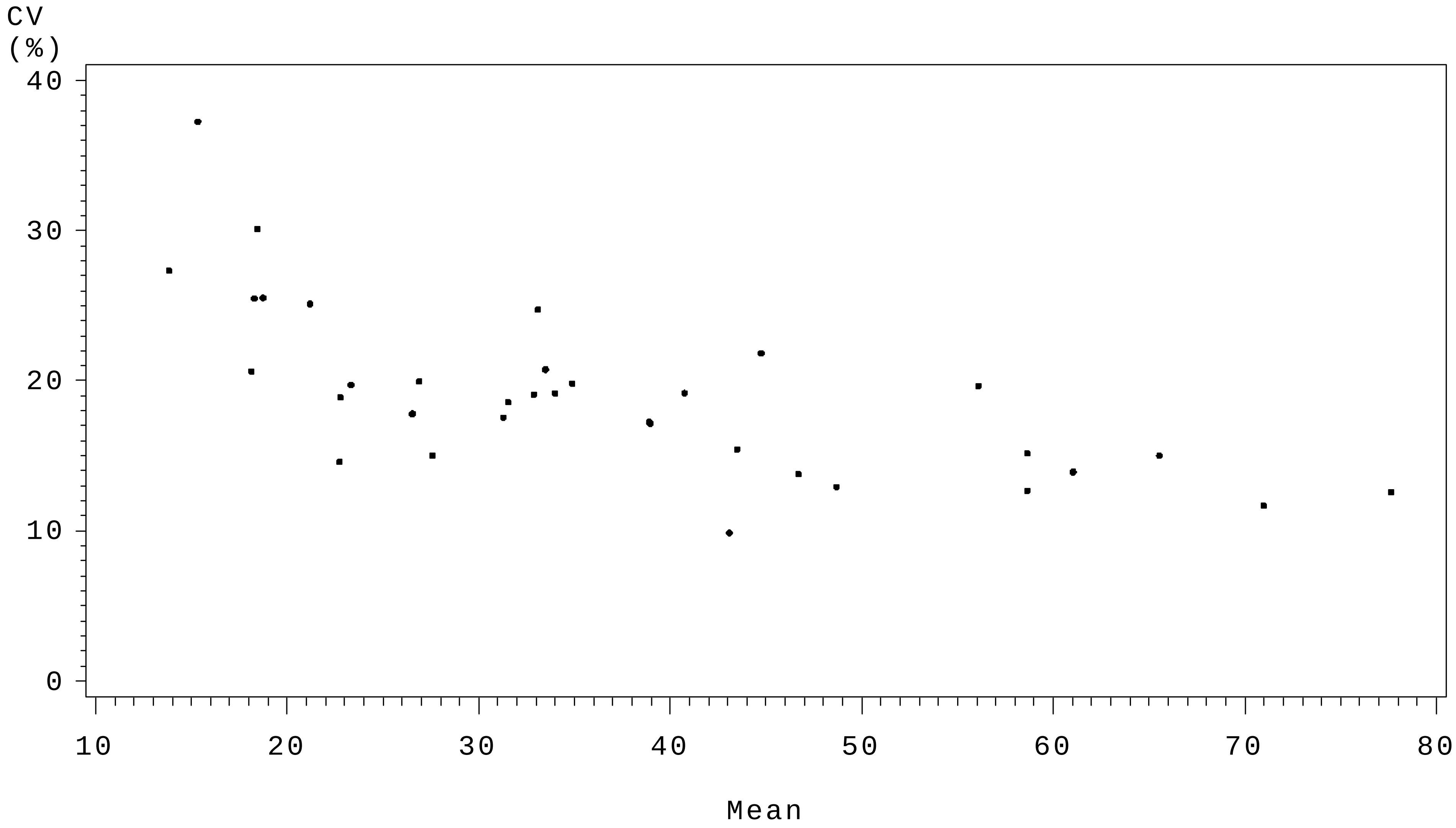
- Research directions : extra sources of variability + confidence intervals

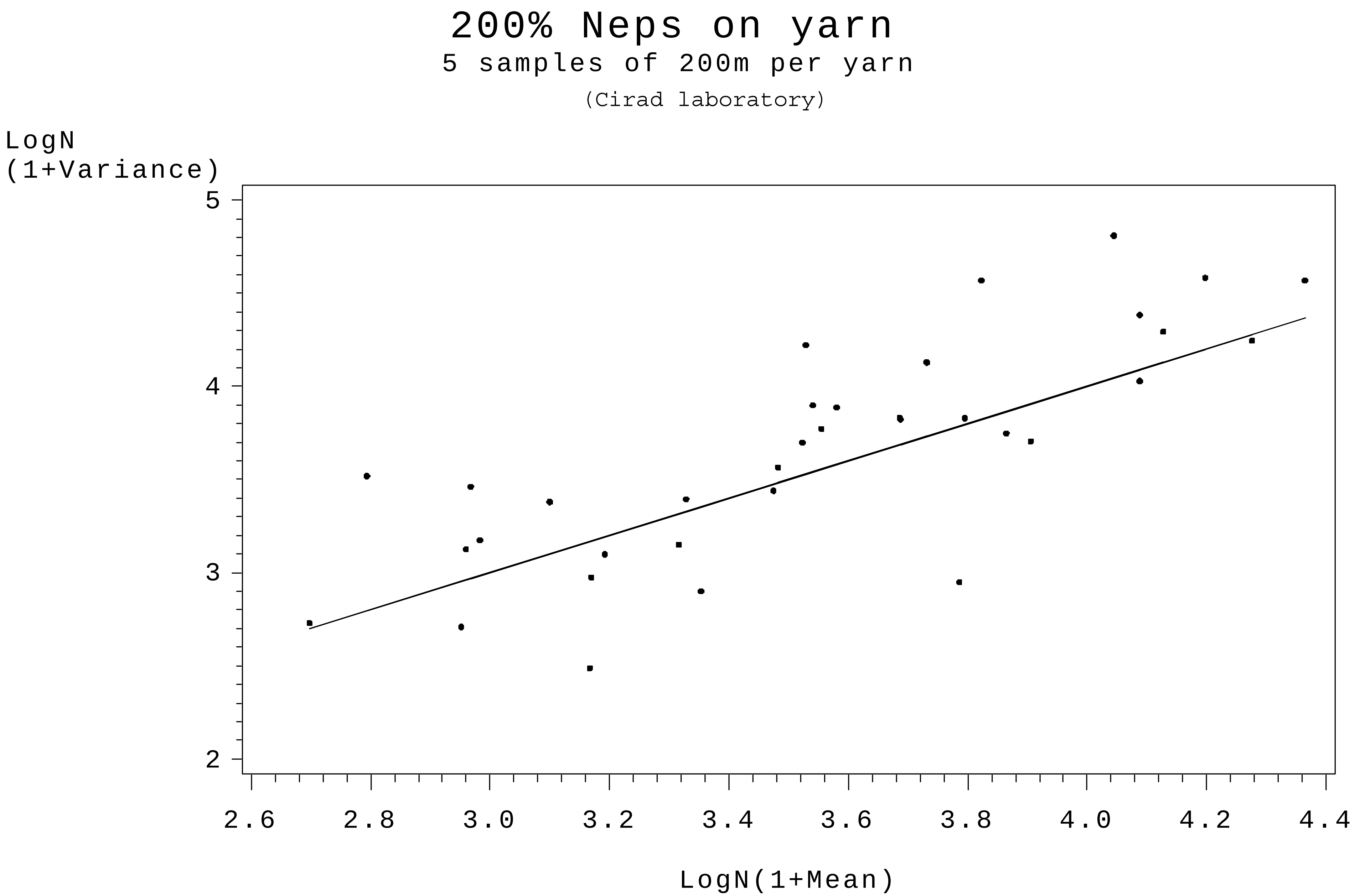

200% Neps on yarn
5 samples of 200m per yarn
(Cirad laboratory)
Standard Deviation
12
11
10
9
8
7
6
5
4
3
10
20
30
40
50
60
70
80
Mean

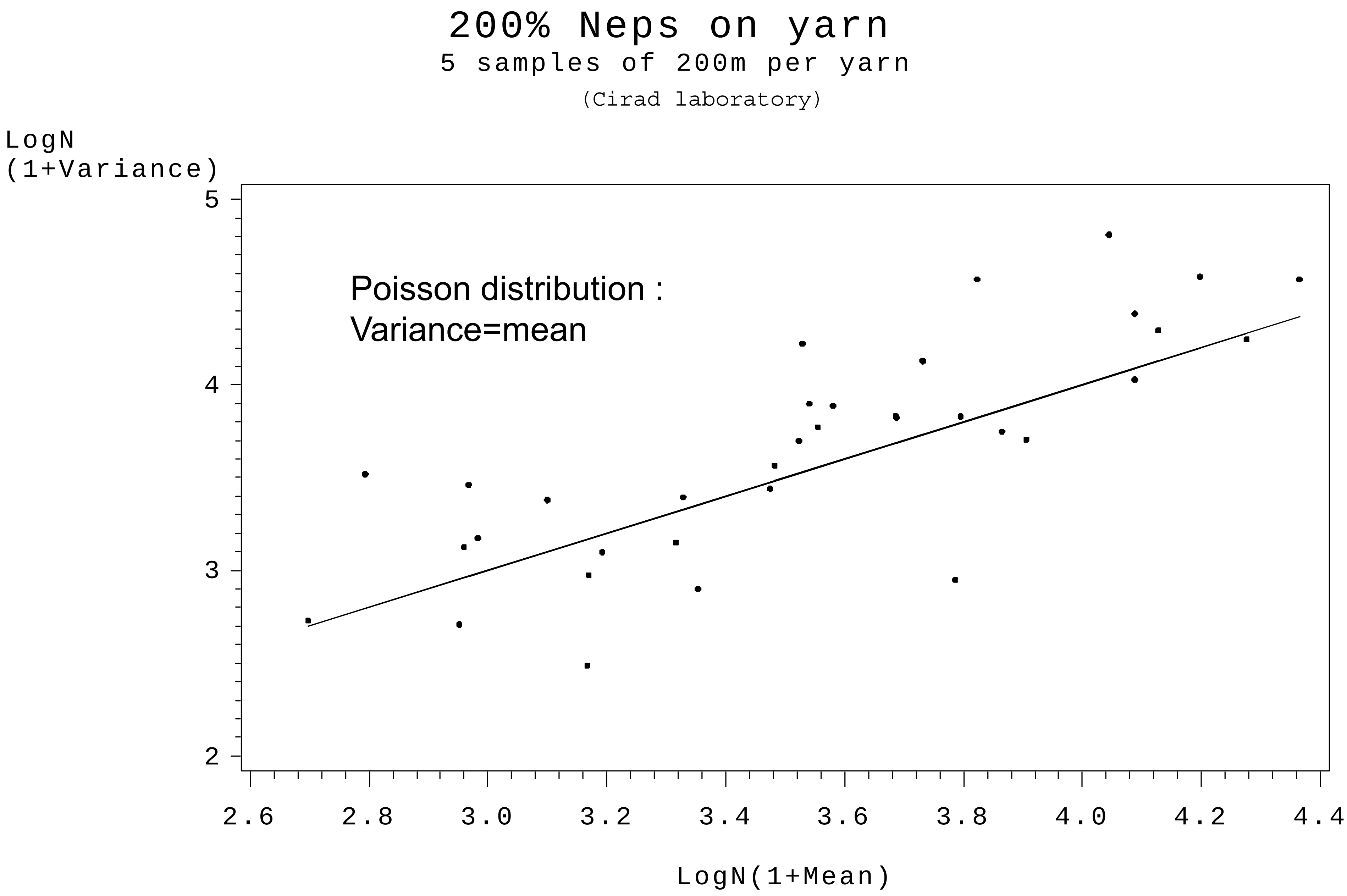
200% Neps on yarn

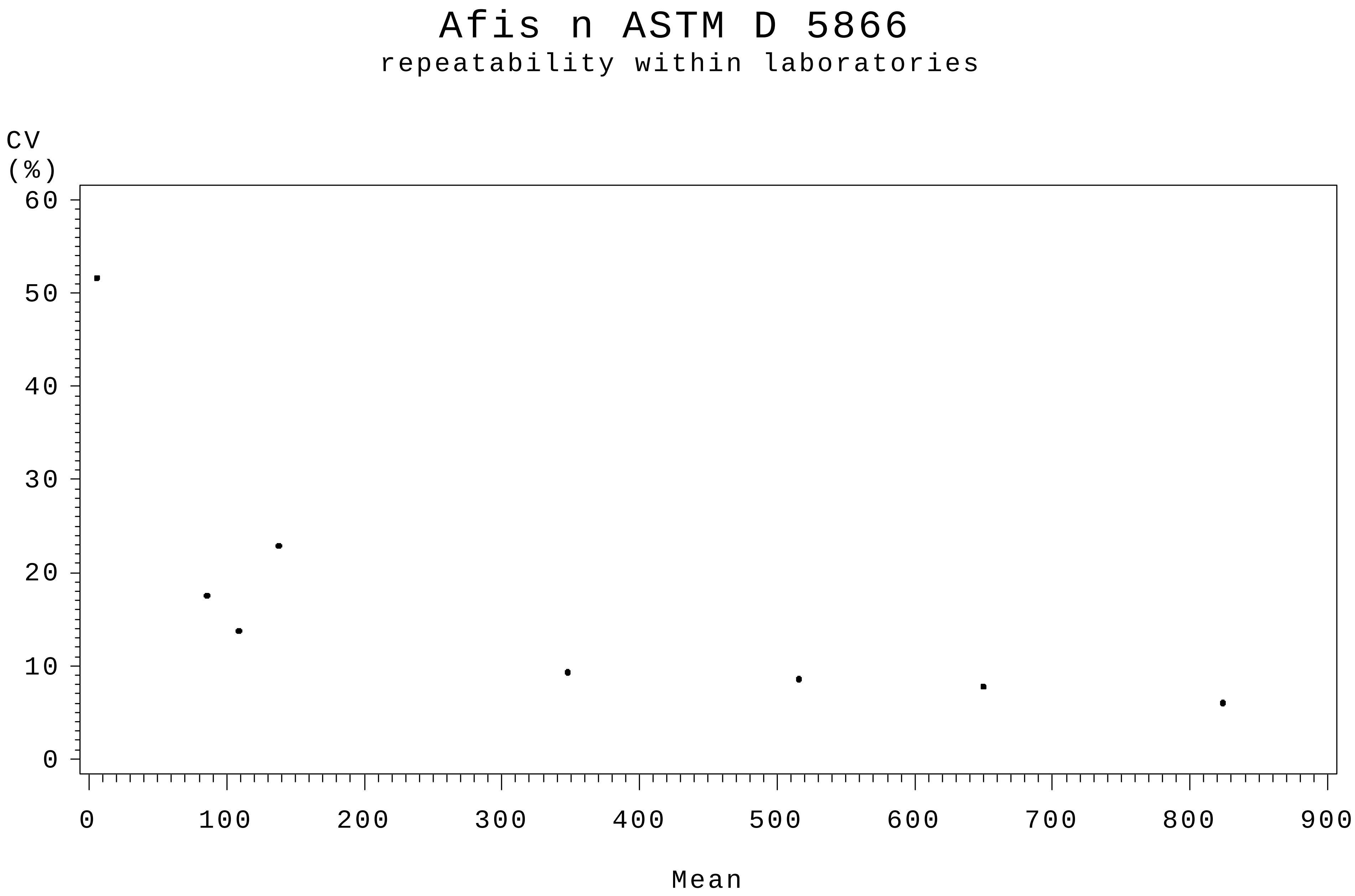
5 samples of 200m per yarn

(Cirad laboratory)

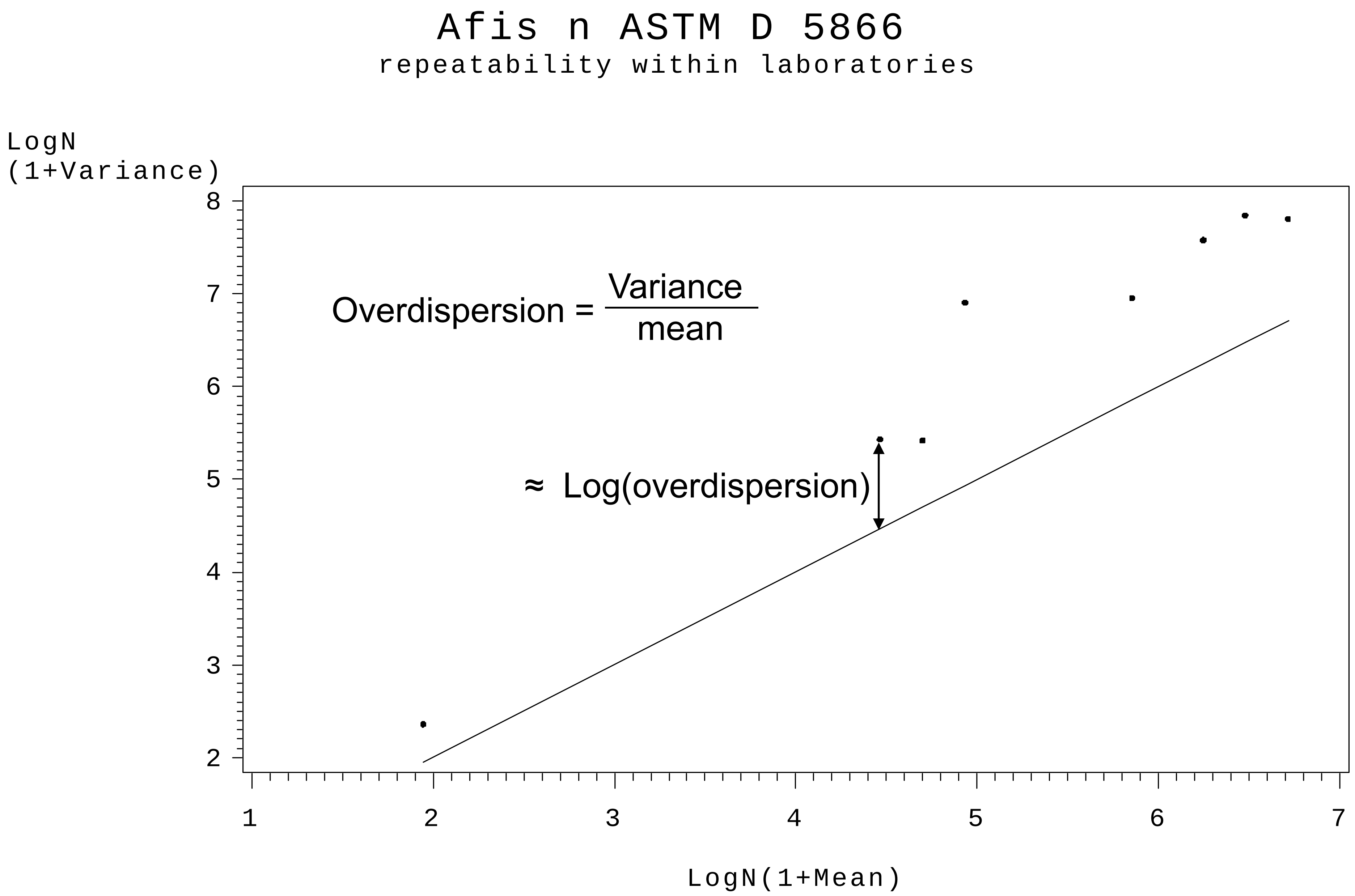



200% Neps on yarn
5 samples of 200m per yarn
(Cirad laboratory)
LogN
(1+Variance)
5
4
3
2
2.6
2.8
3.0
3.2
3.4
3.6
3.8
4.0
4.2
4.4
LogN(1+Mean)


200% Neps on yarn
5 samples of 200m per yarn
(Cirad laboratory)
LogN
(1+Variance)
5
4
3
2
Poisson distribution :
Variance=mean
2.6
2.8
3.0
3.2
3.4
3.6
3.8
4.0
4.2
4.4
LogN(1+Mean)

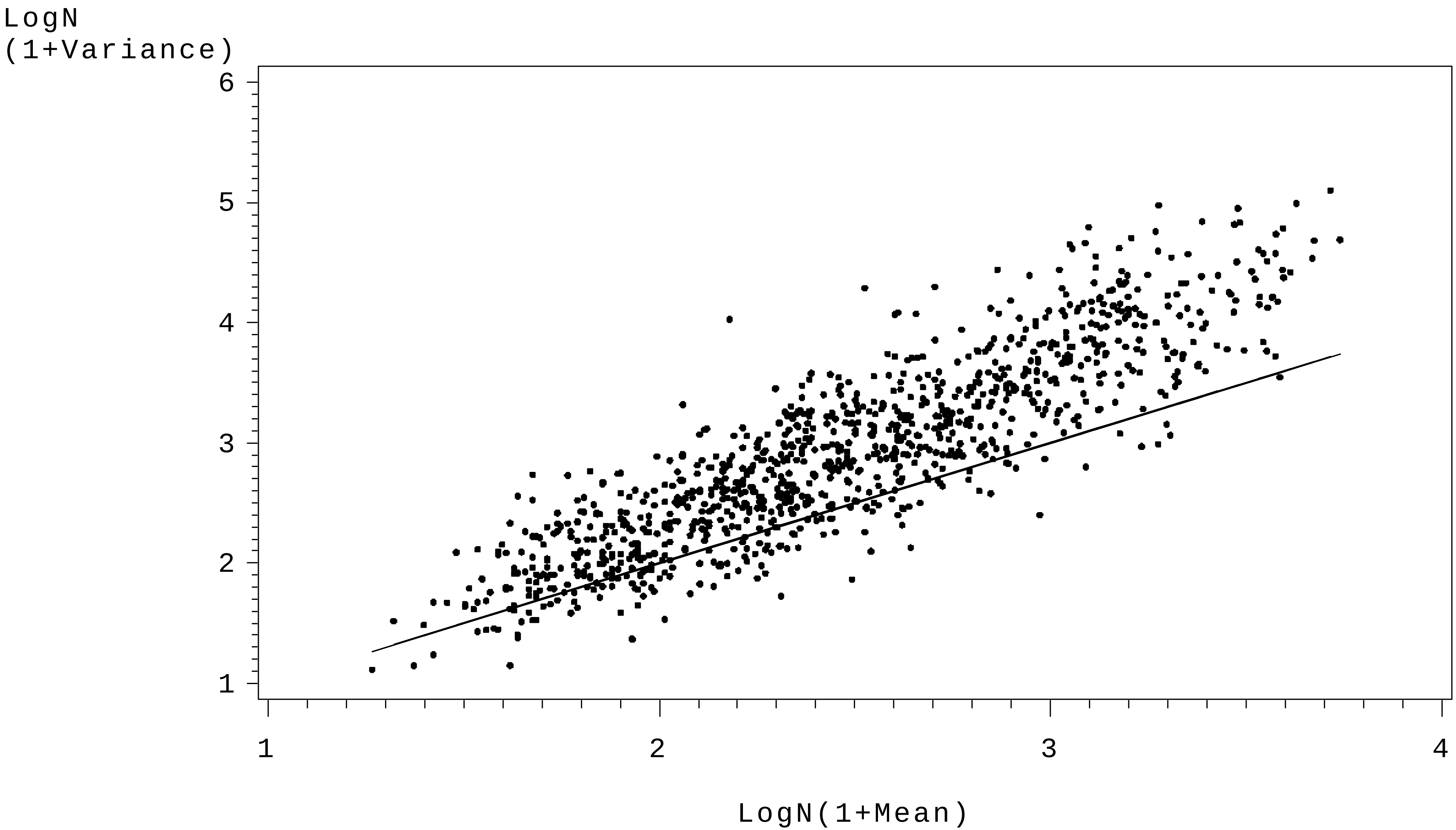


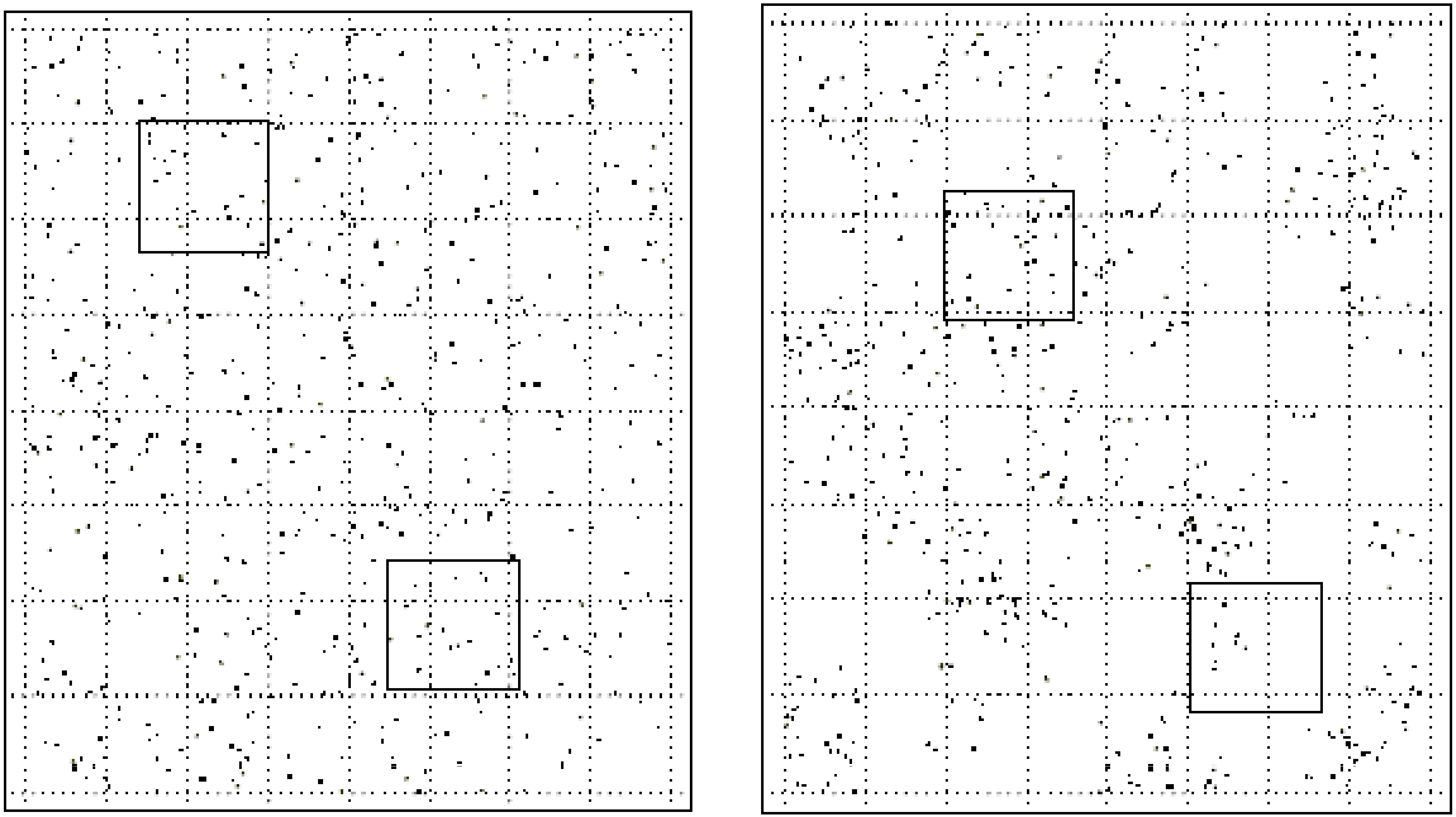
Afis n ASTM D 5866
repeatability within laboratories
CV
(%)
60
50
40
30
20
10
0
0
100
200
300
400
500
600
700
800
900
Mean



Afis n ASTM D 5866
repeatability within laboratories
LogN
(1+Variance)
8
7
6
5
4
3
2
Overdispersion = Variance / mean
≈ Log(overdispersion)
1
2
3
4
5
6
7
LogN(1+Mean)

Trash count
(Cirad laboratory)





(a) (b)

Landmark probability distributions

- Independently located defects in a homogeneous material

  Poisson distribution $\sigma^2 = \mu$

- Patchy located defects in a homogeneous material

  Neyman type A distribution $\sigma^2 = \mu(1+\varphi)$

- Independantly located defects in an heterogeneous material : density varies randomly : compound distribution

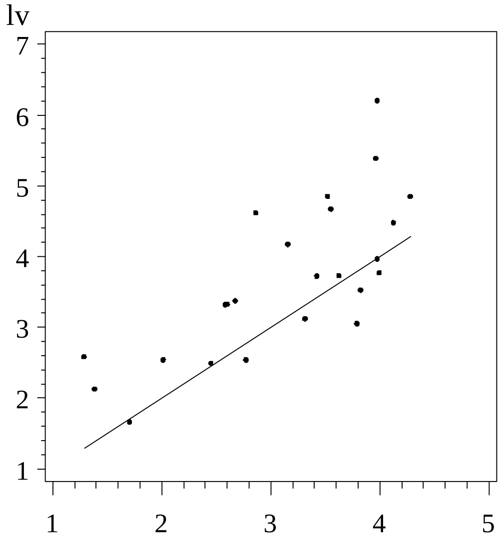
  - Log-normal density :
    Poisson-lognormal distribution $\sigma^2 = \mu\,(1+\mu\varphi)$
  - Gamma density
    Negative binomial distribution $\sigma^2 = \mu(1+\mu/k)$

- Distributions are only landmarks
- Addition of multiple effects : operator, laboratory, calibration yields a more complex compound distribution
- Any of the observed mean-to-variance relationships could be fitted with an overdispersed negative binomial, where

$$\sigma^2/\mu = ø\ (1+\mu/k)$$

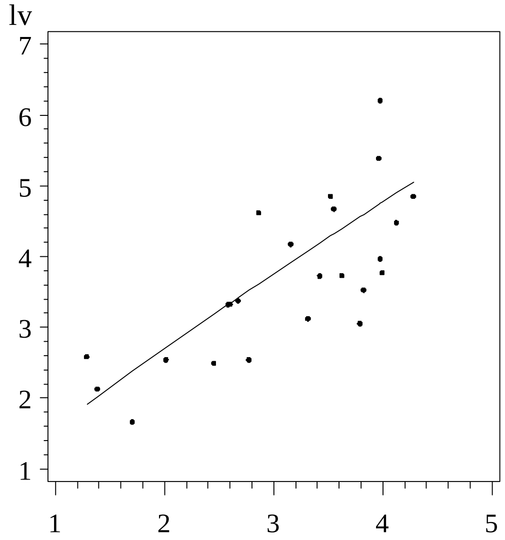
Poisson ϕ = 1

etat=mixed Appareil=H2SD



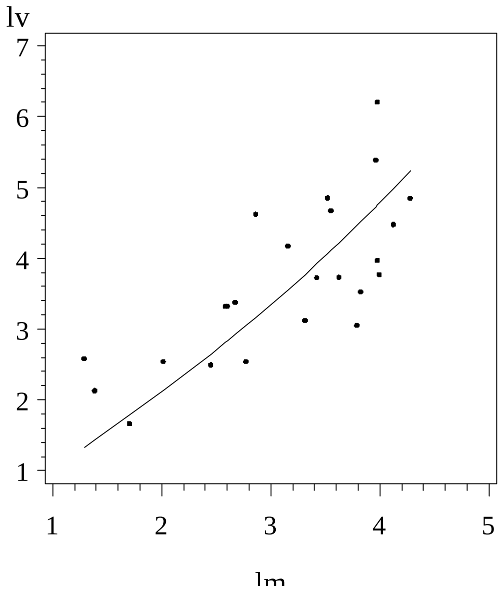
Poisson ϕ = 2.19

etat=mixed Appareil=H2SD



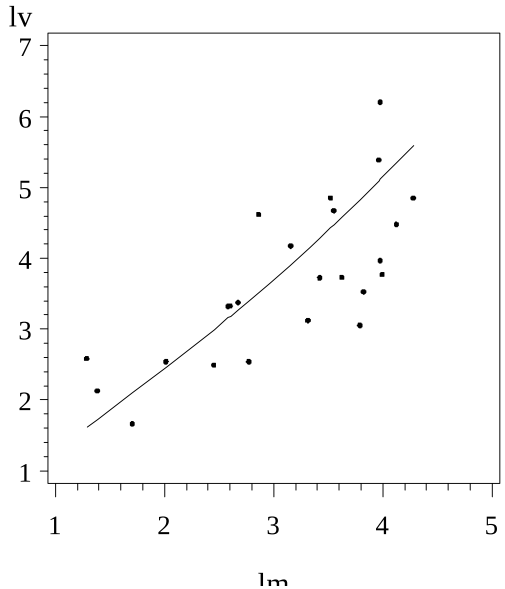
Neg. binomial 1/k=0.0228

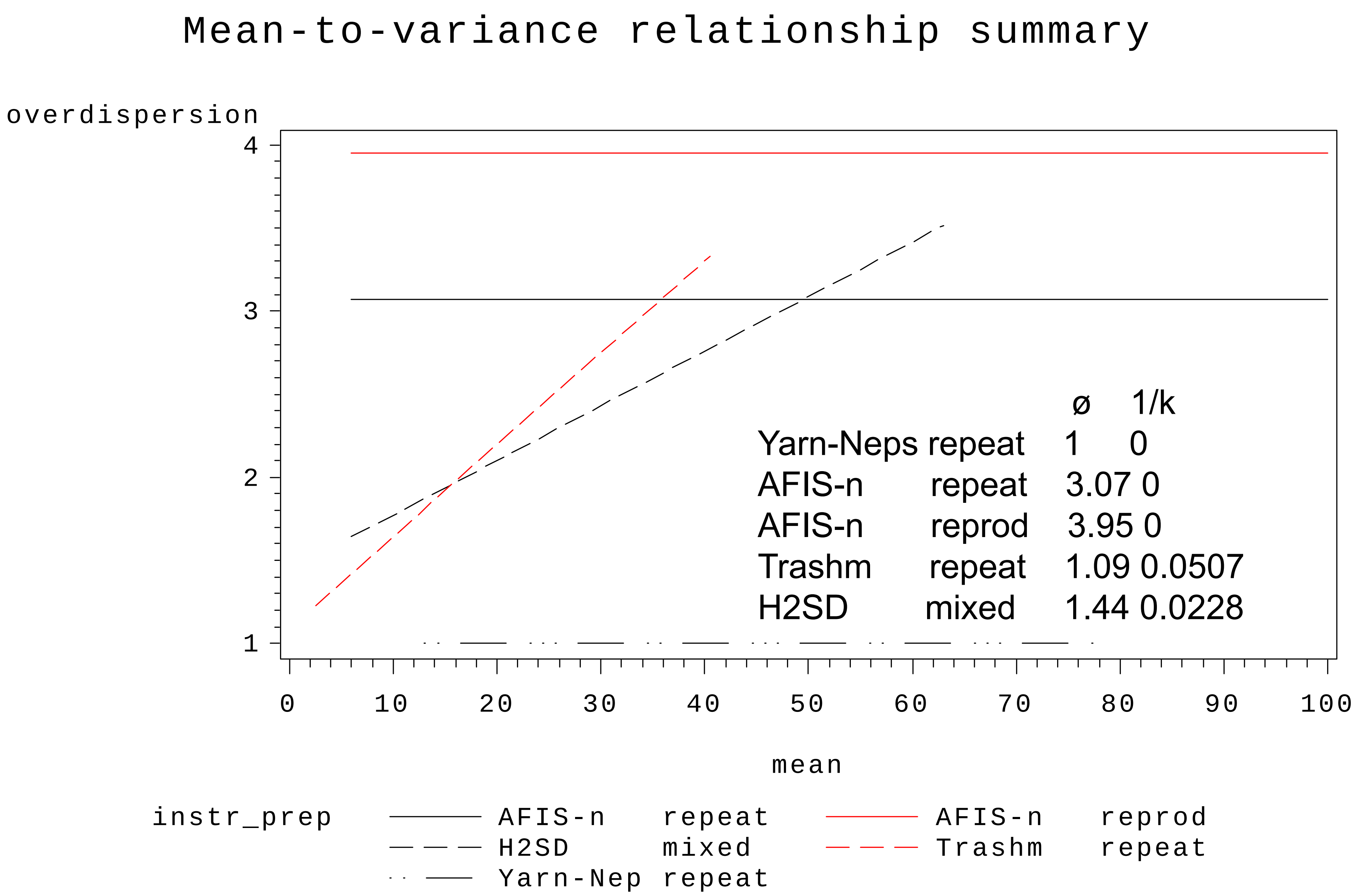
etat=mixed Appareil=H2SD



Neg. binomial 1/k=0.0228 ϕ = 1.44

etat=mixed Appareil=H2SD




Mean-to-variance relationship summary
overdispersion
4
3
2
1
0
10
20
30
40
50
60
70
80
90
100
mean
ø 1/k
Yarn-Neps repeat 1 0
AFIS-n repeat 3.07 0
AFIS-n reprod 3.95 0
Trashm repeat 1.09 0.0507
H2SD mixed 1.44 0.0228
instr_prep
AFIS-n repeat
AFIS-n reprod
H2SD mixed
Trashm repeat
Yarn-Nep repeat

Outline

- Variability of some defects counts on yarn and fiber : evidence and characterization of a mean to variance relationship

- How to analyze calibration and round test experiments

- Research directions : extra sources of variability + confidence intervals

Calibration and round test experiments

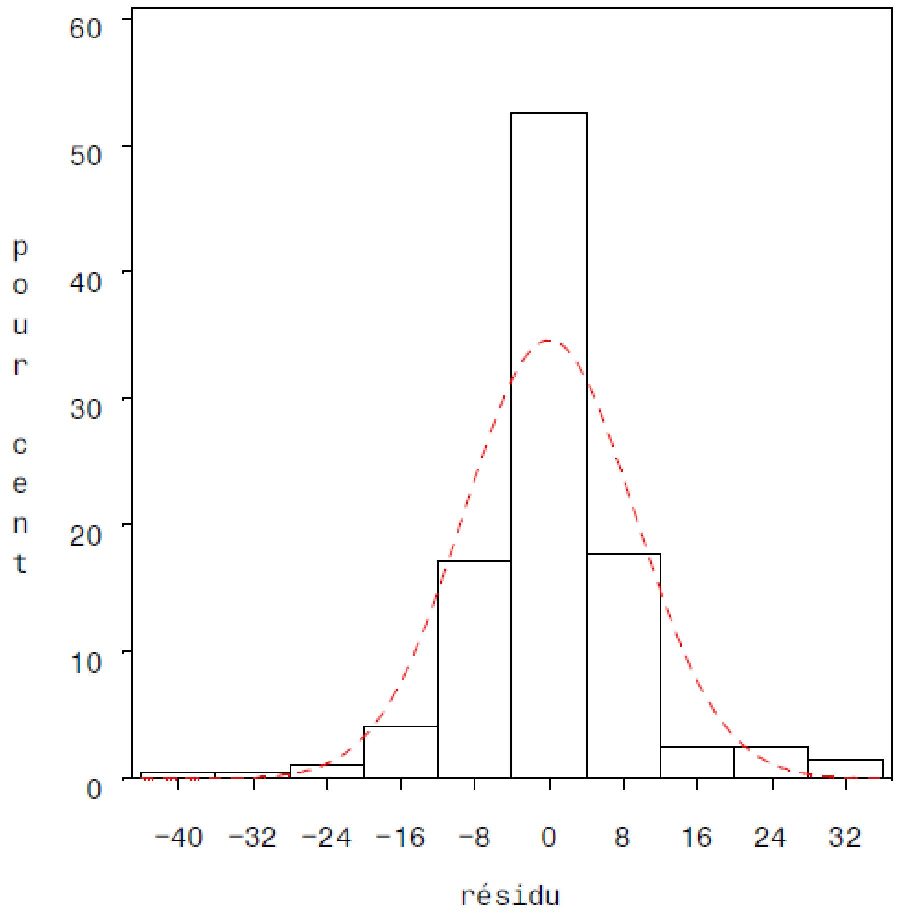
- Regression and analysis of variance are not appropriate for count data, since the distribution of errors are not “normal” (gaussian), and worse not of constant variance.
- This does not always appear clearly, as diagnostics are not displayed as standard in most softwares and spreadsheet toolkits; and also because the measurements conditions are not variable enough to display theses defects clearly.
- We present here an experiment where the measuring conditions have been set on purpose in a way they strongly influence the results of a counting device : this caricatural situation will make more clear the need for alternative methods

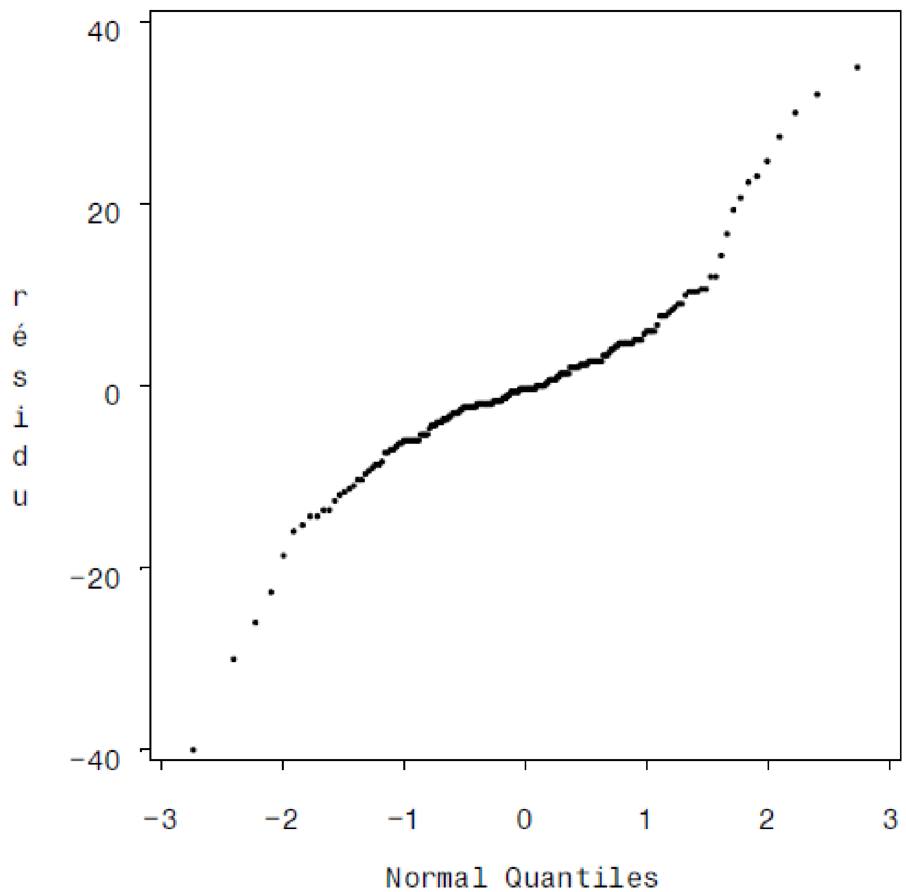
	35%+Enceinte			35%			45%			55%			65%			75%		
	répétition			répétition			répétition			répétition			répétition			répétition		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
coton																		
1	9	17	15	2	8	12	7	8	11	18	26	20	14	11	28	8	11	3
2	24	18	24	13	7	7	14	16	25	33	23	25	29	21	29	8	19	3
3	73	134	90	42	28	33	72	56	73	106	93	106	80	80	121	35	20	61
4	21	32	12	11	11	12	18	12	24	24	30	19	28	33	24	3	5	8
5	140	200	170	59	131	107	140	110	101	149	147	179	172	188	146	66	46	72
6	49	68	37	21	25	14	36	33	48	70	72	60	54	65	81	21	36	28
7	21	20	28	10	11	10	15	21	9	25	26	29	28	28	33	9	14	11
8	53	39	37	13	18	17	31	29	37	40	28	52	39	40	34	21	21	17
9	51	90	55	12	8	23	28	35	28	52	43	60	42	64	55	29	36	28
10	3	5	1	4	1	1	6	3	8	3	3	5	6	5	9	2	3	2
11	38	44	38	24	24	16	18	36	18	31	29	33	30	49	36	25	12	30

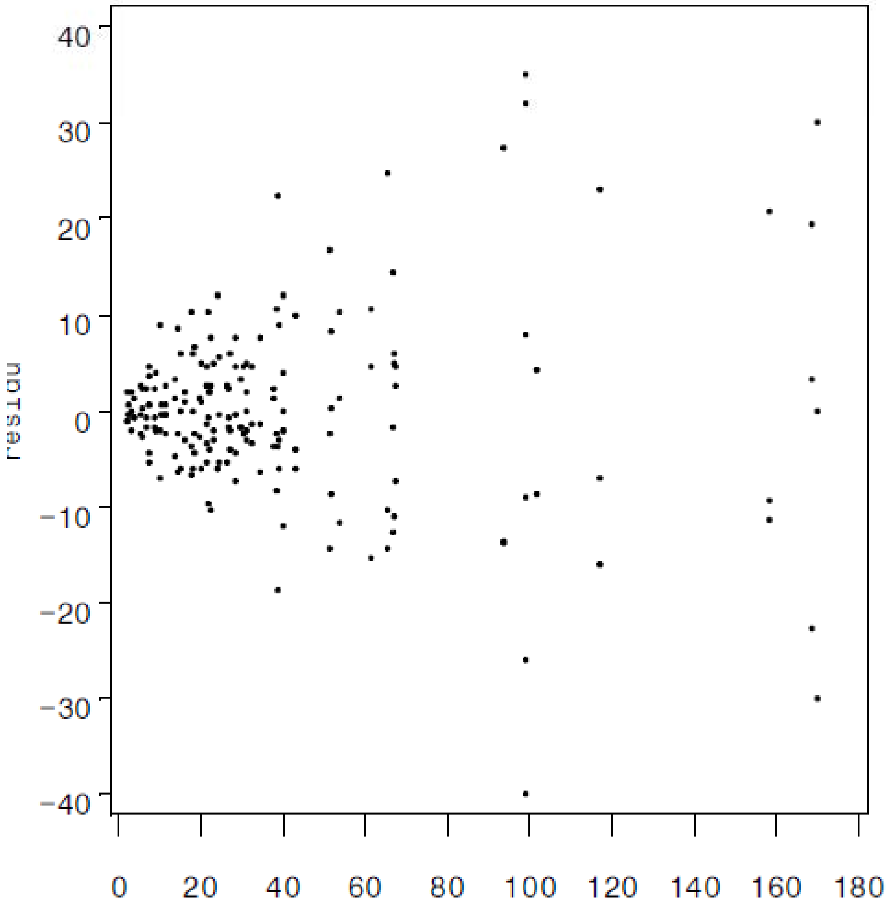
Model for a two-way Anova:

$$Y_{ijk} = m + a_i + b_j + (ab)_{ij} + E_{ijk}$$

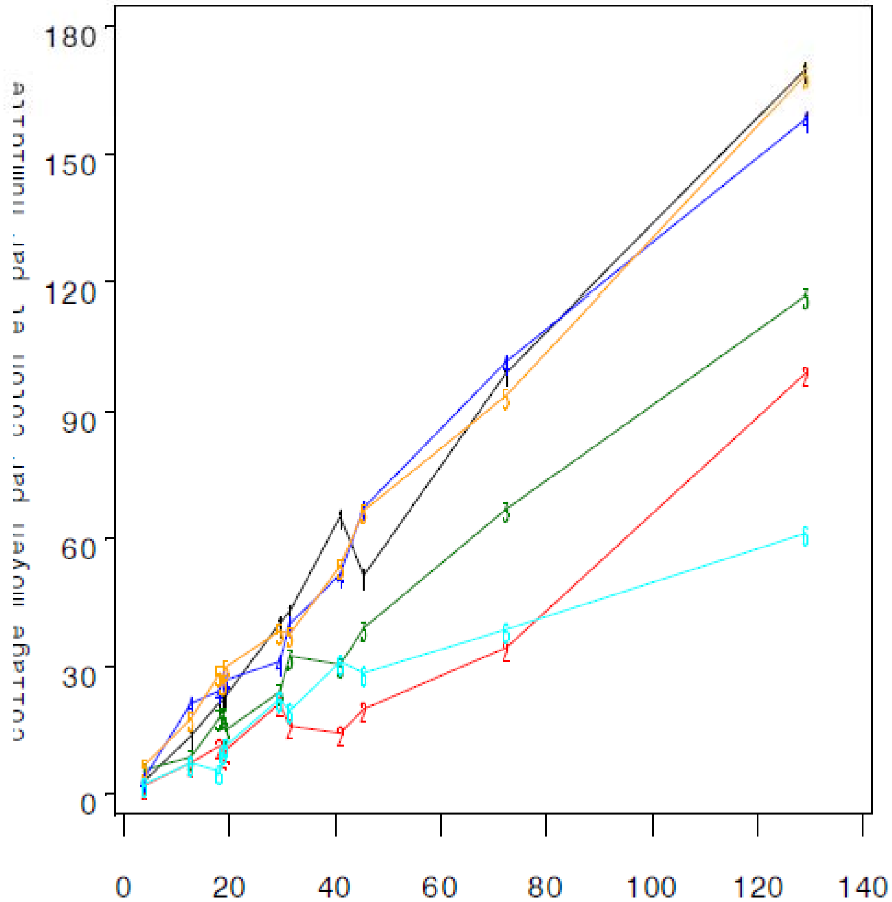
Y_{ijk}= # of SCT sticky spots
i = cotton
j = measurement conditions
k = replicate







Mean to variance relationship



Umbrella-like interaction plot :
Cotton and conditions effects
are multiplicative

Two-way Anova : evidence of a
cotton x measurement conditions interaction

Non-additive effects of measurement conditions
=> cannot propose a calibration coefficient
valid for all cottons

Analysis of variance on the SCT # sticky dots, cubic root transformed to stabilize variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	65	194.2782433	2.9888961	35.11	<.0001
Error	132	11.2378639	0.0851353		
Corrected Total	197	205.5161072			

R-Square	Coeff Var	Root MSE	rac3_col Mean
0.945319	9.573093	0.291780	3.047913

Source	DF	Type I SS	Mean Square	F Value	Pr > F
HUM	5	30.5636193	6.1127239	71.80	<.0001
coton	10	156.3502085	15.6350208	183.65	<.0001
HUM*coton	50	7.3644155	0.1472883	1.73	**0.0072**

Generalized linear model replaces analysis of variance :

$$\mathrm{Log}(\ \mu_{ijk}\ ) = m + ai + bj + (ab)ij + Eijk$$

$$E[Yijk] \quad = \mu$$

$$V(Yijk\ ) \quad = \mu \quad \text{(Poisson distribution)}$$

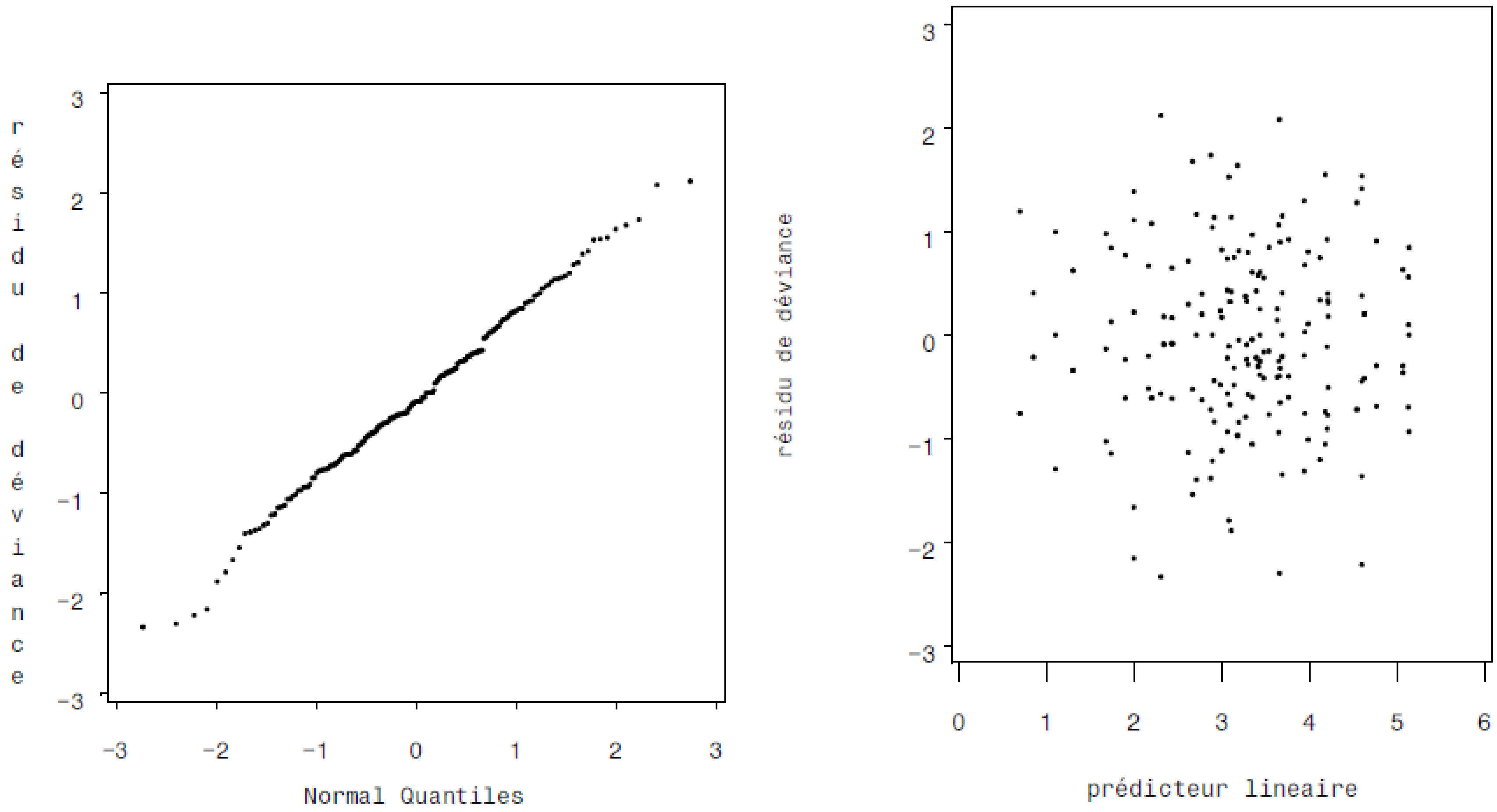
or $$V(Yijk\ ) \quad = \mu\ ø\ (1+\mu/k) \text{ (Overdispersed distribution)}$$

No more interactions : calibrations coefficients are now valid regardless of the particular cotton being tested

LR Statistics For Type 1 Analysis

Source	2*Log Likelihood	Num DF	Den DF	F Value	Pr > F	Chi-Square	Pr > ChiSq
Intercept	45598.5488						
HUM	45632.0141	5	132	4.67	0.0006	23.35	0.0003
coton	46097.8166	10	132	32.50	<.0001	324.99	<.0001
HUM*coton	46174.2179	50	132	1.07	0.3789	53.30	0.3483

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits	
Intercept		1	2.8593	0.0845	2.6937	3.0249
HUM	35%	1	-0.0438	0.0802	-0.2009	0.1133
HUM	35%+E	1	0.7968	0.0732	0.6534	0.9402
HUM	45%	1	0.4297	0.0758	0.2811	0.5782
HUM	55%	1	0.8308	0.0731	0.6875	0.9741
HUM	65%	1	0.8640	0.0730	0.7209	1.0072
HUM	75%	0	0.0000	0.0000	0.0000	0.0000
coton	1	1	-0.8721	0.1083	-1.0844	-0.6598
coton	2	1	-0.4792	0.1004	-0.6760	-0.2823
coton	3	1	0.8635	0.0867	0.6935	1.0335
coton	4	1	-0.5167	0.1010	-0.7147	-0.3186
coton	5	1	1.4698	0.0843	1.3045	1.6351
coton	6	1	0.4006	0.0898	0.2245	0.5766



résidu de déviance
Normal Quantiles
-3
-2
-1
0
1
2
3
résidu de déviance
prédicteur lineaire
0
1
2
3
4
5
6

The mean-to variance relationship is enough
to provide means of analysis
of round tests and calibration data
with a generalized linear model,
where usual regression and Anova fail.

The results are sound and intelligible

Though now standard in medical and insurance
research, the generalized linear model
is not part of the toolkit
of the engineers and university graduates,
(except when specialized in Statistics).

Outline

- Variability of some defects counts on yarn and fiber : evidence and characterization of a mean to variance relationship
- How to analyze calibration and round test experiments
- confidence intervals + research directions on extra sources of variability

The mean-to variance relationship is not enough to provide means of calculating exact confidence intervals and litigation risks. Approximate results are available only when the count numbers are high enough.

Checks of distributions have to be made on cottons, especially in the region of maximum litigation risks.

compound Poisson distribution,

$$P_r = \int_0^\infty \frac{\lambda^r e^{-\lambda}}{r!} f(\lambda)\, d\lambda, \qquad r = 0, 1, 2 \cdots$$

The generalized linear model can be applied to more complicated designs to allow the estimation of different flavours of reproducibility :
changing week, month, technician, device, labs, …

Different scales of observation (specimen, sample, bale, module…) mean different sources of variability adding up. The resulting distribution can be seen as a continous mixture of Poisson distribution of various expectations, and the resulting distribution can be calculated by integration.

Funded by CSITC project, some investigations are in progress to investigate variability within the coton production in Africa, I hope my African colleagues and I will show you the results in two years time.

Measurements based on counts : variability and methods of analysis

Useful measurement must be

1 correlated to practical properties of the analyzed material

2 reproducible, i.e. its reproducibility has been measured with proper indexes.

Apparent paradoxes displayed by classical variability indexes such as standard deviation or CV on counts can be overcome by considering a simple relation between mean and variance.

This leads to a general way of measuring the variability, from repeatability to different flavours of reproducibility.

Is there a threshold beyond which an instrument
should be rejected ?

A high overdispersion does not imply the instrument is not reliable (see e.g. AFIS-n). However the customer should be informed about the number of replicates needed to achieve a given precision. Commercial norms for the number of replicates should be derived from these numbers.

The customer should also be informed of the procedure to calibrate his instrument and sets of standard cottons should be available to do so.

Trash area
(Cirad laboratory)